Research on Pilots ' Mental Workload Classification in Simulated Flight

Jinna Xue*
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, China
E-mail:xgenia@163.com
*Corresponding author

Changyuan Wang
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, China
E-mail: Cyw901@163.com

***Abstract*—The problem of human-computer interaction mental workload in flight driving has great reference value for the prevention of safety hazards in aviation driving. This paper analyzes and studies the classification method of mental workload in flight driving by designing different simulated flight experiment tasks. This study uses a combination of EEG signals and subjective evaluation, through the use of convolutional neural networks and long short-term memory network method of combining EEG signals for research and analysis. The accuracy of EEG signal classification is as high as 94.9 %. NASA-TLX evaluation results show that there is a positive correlation between task load difficulty and evaluation score. The results show that the combination of convolutional neural network and long short-term memory network is suitable for pilots ' mental workload classification. This study has important practical significance for flight accidents caused by pilots ' mental workload.**

Keywords-Mental Workload; EEG; Convolutional Neural Network; Long Short-Term Memory Network; Subjective Evaluation Method

I. INTRODUCTION

With the emergence and rapid development of the Internet, people's load on cognitive resources has also increased, and the demand for mental workload optimization has become increasingly prominent [1]. Studies have shown that too high and too low cognitive load will have a negative impact on job performance [2]. Pilots need to obtain more instrument data information during driving, and combine it with the working environment for comprehensive analysis, so as to make corresponding decisions quickly and accurately [3]. In the relevant operation process requires a high degree of concentration, may produce large mental labor [4], will not only affect the work efficiency, and even cause serious accidents. Thus, in human-computer interaction, mental labor gradually occupy the advantage. The intensity of mental labor can be measured by mental workload, which is mainly generated in projects where operators use limited mental resources for task processing [5]. Appropriate mental workload helps the operator to maintain good task performance, while too high or too low mental workload will adversely affect the performance in the operation.

According to the aviation accident research report, 60 % to 90 % of flight accidents occur in flights where the pilot 's mental workload intensity is too high and the stress task level is high [6]. Complex human-machine interaction operations such as air defense missiles, medical rescue and efficient driving are often accompanied by high-load human-machine interaction tasks [7]. For example, in the operation of missile weapons, it is necessary to achieve accurate and fast analysis of space intelligence such as the course and speed of combat targets, and to quickly realize the assessment of the threat of combat targets for the implementation of complex actions such as target tracking and attack [8-10]. Because there is a large degree of human-computer interaction tasks in this process. The high-intensity operation process can easily cause the mental operator to fall into an overload state, which greatly reduces the safety and reliability of the operator 's operation process and easily causes safety accidents. Therefore, the study of mental workload is of great significance to the safety and reliability of human-computer interaction.

The concept of mental workload was first proposed in the 1980 s. Scholars have mainly conducted in-depth research and discussion on the causes, internal mechanisms, and measurement methods of mental workload [11]. In the past forty years, researchers have continued to pay attention to the mental workload of operators. The physiological measurement method has greater advantages in terms of sensitivity to changes in brain load and digital quantitative analysis, but there are greater difficulties in terms of post-processing of data and other aspects, a certain knowledge base is required for data analysis, and the highly sensitive measurement method is susceptible to interference from external factors. In response to these difficulties, some scholars have proposed the use of eye-movement metrics for the assessment of brain load [12-14]. Liu Yan [15] et al. conducted a correlation analysis of pilots' brain load in terms of typical issues such as fatigue and sleep, and obtained from the experimental results that the correlation degree coefficient between brain load and multidimensional fatigue measurements was around 0.5, while the correlation degree coefficient between brain load and Pittsburgh sleep quality was as high as 0.59. From the results of this study, it can be obtained that the degree of pilots' brain load is influenced by the degree of fatigue and sleep. Wang Lei et al. from Civil Aviation University of China [16] conducted a study on the brain load characteristics of pilots based on flight task context routes, in which they used different task difficulties to carry out relevant studies on the brain load characteristics of pilots. The experimental results are of great reference value for the study of pilot brain load classification methods. In this paper, we will draw on the relevant pilot brain load characteristics to investigate pilot brain load classification methods.

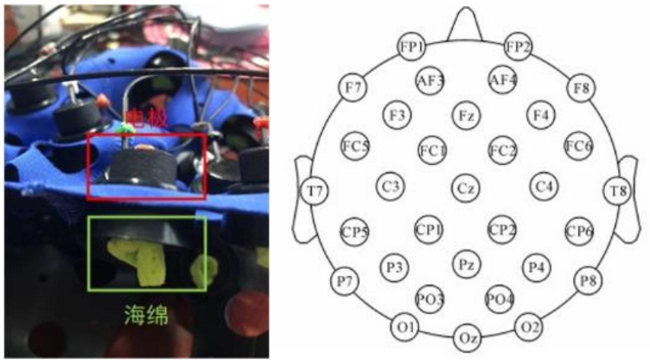
II. EXPERIMENT

A. Experimental preparation

The experiment was conducted under a simulated flight platform and eight male graduate student volunteers aged 22-30 years were recruited as subjects for data collection. The subjects were all in good health, had normal or corrected vision, were right-handed and had good sleep conditions. The subjects were familiarised with the details of the experimental operation and related precautions before the simulation. When conducting the experiments, only the subjects were left alone for each trial in order to avoid interference from the outside environment.

The equipment for this pilot mental workload classification study contains a DELL computer, three high definition displays (2560 x 1440 resolution), EEG signal synchronisation acquisition equipment and the test deployment platform is shown in Figure 1. The flight simulation platform contains a monitor, flight joystick and mainframe. The flight simulation was conducted using the DCS World digital battlefield game, using a free to fly Su-25T fighter aircraft designed and manufactured by Sukhoi in Russia.

Figure 1. Simulated flight environment experimental platform



(a) EEG signal acquisition device (b) EEG cap port diagram

Figure 2. ErgoLAB EEG device

The experimental equipment used in this experiment was the ErgoLAB human-computer environment synchronous platform EEG measurement system provided by Beijing Jinfa Technology Co. The system adopts the international standard 10-20 standard lead semi-dry electrohydraulic stage with an output bit 32-lead EEG signal; a wet sponge is used as the conduction medium, as shown in Figure 2(a).

B. Experiment design

In this simulated flight experiment task, each subject performed three main simulated flight experiments containing three mental-load flight tasks: control, low load and high load. The cockpit dial information that needed to be monitored during the simulated flight corresponding to the different difficulty tasks during the simulated flight was shown in Table 1. The flight information contained in the cockpit interface [17] includes indication of airspeed table, pitch scale information, no altimeter information, heading angle, roll angle, direction pod, landing gear status and engine status. During a flight task, the subject is required to respond accurately and quickly to the information presented in the interface.

TABLE I. INFORMATION ON THE DIALS TO BE MONITORED FOR DIFFERENT DIFFICULTY TASKS

High loads	**Low loads**	**Control group**
Indicated airspeed meter	Indicated airspeed meter	—
Pitching scale information	Pitching scale information	—
Altimeter information	Altimeter information	—
Directional angle information	—	—
Rolling corners information	—	—
Status of the steering compartment	—	—
Landing gear conditions	—	—
Engine status	—	—

C. The experimental process

For different subjects, the experiments are conducted one by one. Each subject is required to perform a control, low and high load flight for each experiment. In the same flight environment, the operator is prompted with different difficulty tasks based on the flight heading status and is required to quickly and accurately monitor the corresponding dial information.

The NASA-TLX Subjective Rating Scale [18] was used in the experiment. The scale of [0,100] was used to indicate the scale's score range to rate the subject's level of mental workload on six dimensions, with higher scores indicating higher levels of workload.

D. Data recording

During the experiment, simultaneous acquisition of EEG signals was performed using EEG acquisition equipment. At the end of each set of simulated flight experiments, subjects are required to complete the NASA-TLX measurement form, which is used to evaluate the subjective load of the subjects.

III. METHODS

In this study, a modified long and short term memory network was used to classify the EEG signals for mental workload research, and the flight mental workload was studied and analysed by combining subjective evaluation methods.

A. Data pre-processing

During the experimental process, the hardware equipment will inevitably be disturbed by external factors. Therefore, the raw EEG signal needs to be pre-processed before the mental workload analysis is carried out, and the processing flow is shown in Figure 3.

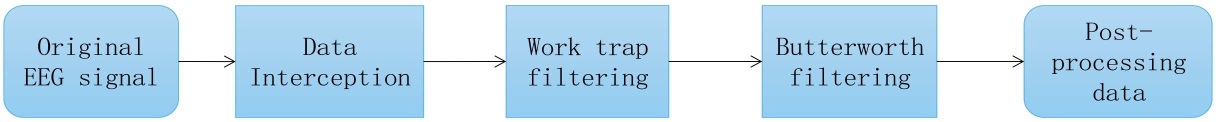


Figure 3. Figure. 3 Flow chart of data processing

Related studies have shown that most EEG signals are concentrated between the range of 0.05-50Hz, while event-related desynchronisation and event-related synchronisation (ERD/ERS) patterns are mainly manifested in μ rhythms (8-13Hz) and β beats (14-30Hz), so this study used Butterworth filters for 50Hz IDF trapping and 0.05-30Hz bandpass filtering, and Butter The Butterworth band-pass filtering removed the fine burrs of interference from the original data [19]. A comparison of the data before and after processing is shown in Figure 4.

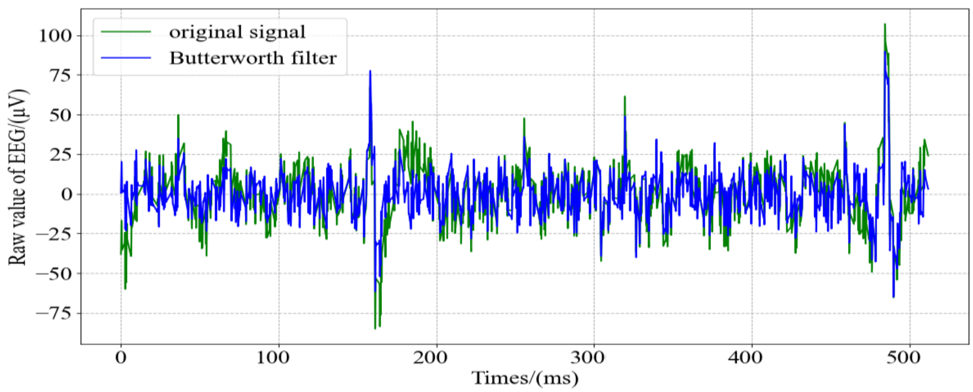


Figure 4. Comparison of EEG waveforms before and after data processing

B. EEG features extraction

According to recent research in cognition, psychology and psychiatry, EEG information about mental activity can be divided into five bands: delta (δ), theta (θ), alpha (α), beta (β) and gamma (γ), and the frequency range of each band is shown in Table 2.

TABLE II. TABLE OF FREQUENCY BANDS OF EEG SIGNALS

Frequency band range(Hz)	Band names
1-4	delta (δ)
4-8	theta (θ)
8-13	alpha (α)
13-30	beta (β)
>30	gamma (γ)

In this paper, the discrete Fourier transform is used for EEG signal feature extraction [20]. The power spectrum was estimated mainly using the Welch method with the power spectral densities of the σ, θ, α and β frequency bands, using a Hamming window with a window size of 256 and a 50% overlap of adjacent window segments. In calculating the power of different frequency bands, the sum of the power density of the band, i.e. the area within the corresponding frequency band of the corresponding power density curve, can be calculated as shown in equation (1):

$$\begin{cases} P_\delta = \int_1^4 \mathrm{psd}(f)df \\ P_\theta = \int_4^8 psd(f)df \\ P_\alpha = \int_8^{13} psd(f)df \\ P_\beta = \int_{13}^{30} psd(f)df \end{cases} \quad (1)$$

Where, $psd(f)$ denotes the power spectral density function and f indicates the frequency of the signal. This paper uses the power under the four fundamental frequency bands spectral densities for mental workload classification judgement.

C. Network model

The network model proposed in this paper is shown in Figure 5. The model is a new input convolutional neural network in front of the long and short-term memory network for local feature extraction of the EEG signal. The extracted local features are then fed into the long and short-term memory network for training. The loss function for model training is L1-loss and its expression is:

$$L1-loss = \sum_{i=1}^{n} \left| x_{true} - x_{predicted} \right| \quad (2)$$

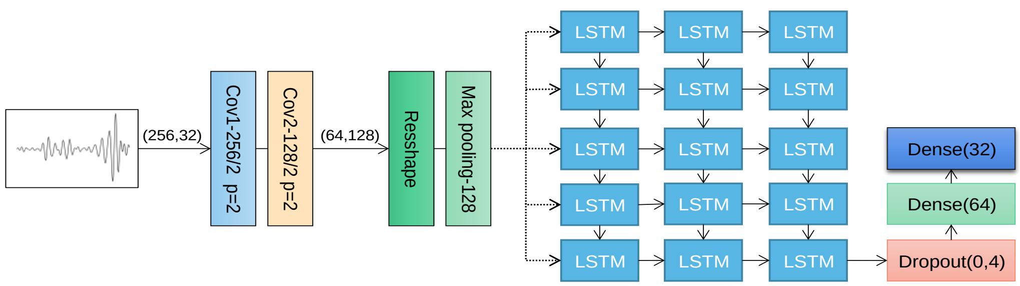


Figure 5. Network model diagram

D. LSTM networks

The Long Short-Term Memory (LSTM) is a recurrent neural network (RNN) [21]. The LSTM contains a three-part structure of input gates, forgetting gates and output gates, which is shown in Figure 6 and mainly outputs (0,1) through the sigmoid activation function values between.

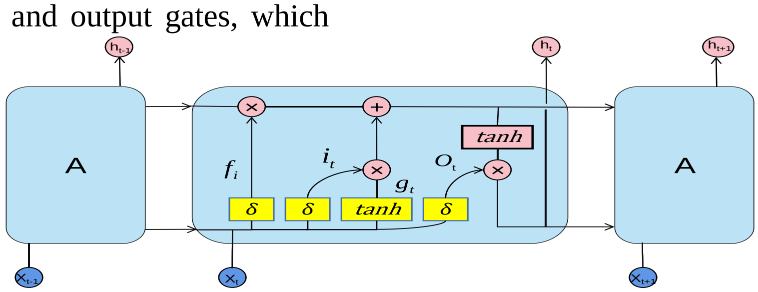


Figure 6. LSTM structure diagram

The role of the forgetting gate is to control the discard rate of the delivered information, as implemented in equation (3).

$$f_t = \sigma(W_f^T \times s_{t-1} + U_f^T \times x_t + b_f) \quad (3)$$

Where δ indicates the activation function, usually set to a sigmoid function, w_f^T indicates the forgetting gate weight matrix, U_f^T is the weight matrix between the input and hidden layers of the forgetting gate, and b_f indicates the bias of the forgetting gate, s_{t-1} is the previous moment's output value and x_t is the current moment's input value. The closer the forgetting gate output value is to 1, the more information is retained, and the closer it is to 0, the less information is retained.

The input gate determines at what percentage of the current moment's input information is fed into the memory information stream, and is calculated similarly to the forgetting gate, as shown in equation (4).

$$i_t = \sigma(W_i^T \times s_{t-1} + U_i^T \times x_t + b_i) \qquad (4)$$

The output gate is mainly used to control the amount of information updated by the next layer of the network and is implemented as shown in equation (5).

$$O_t = \sigma(W_O^T \times s_{t-1} + U_o^T \times x_t + b_o) \qquad (5)$$

$$s_t = O_t \times \tanh\left(C_t\right) \qquad (6)$$

σ is usually a sigmoid function. This structure expands the effectiveness of the action of the current amount of information so that it can both suppress the current information and output it normally through equation (6) combined with the memory information to obtain the output value at the current moment.

IV. RESULTS AND ANALYSIS

Based on the results of the NASA-TLX subjective evaluation table obtained statistically, the sensitive assessment index scores for changes in mission difficulty can be obtained as shown in Figure 7. According to the ANOVA results, the NASA-TLX subjective evaluation scores show an extremely significant difference ($p<0.01$) and become larger as the difficulty of the mission increases and the load level increases.

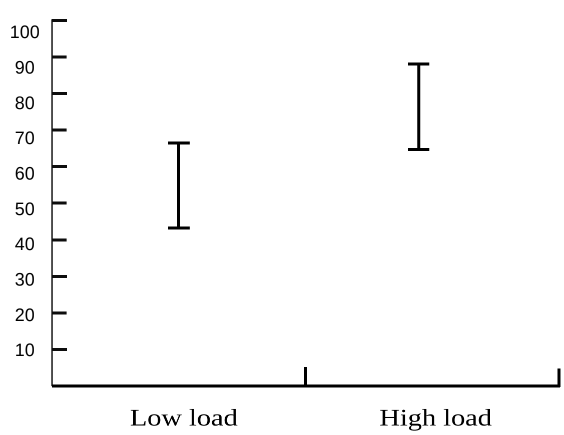


Figure 7. NASA-TLX evaluation results

Based on the results of the NASA-TLX subjective evaluation table obtained statistically, the sensitive assessment index scores for changes in mission difficulty can be obtained as shown in Figure 7. According to the ANOVA results, the NASA-TLX subjective evaluation scores show an extremely significant difference ($p<0.01$) and become larger as the difficulty of the mission increases and the load level increases.

TABLE III. MODEL TRAINING PARAMETERS

Name of the layer	Hyperparameters	Input vector	Output vector
EEG_Input	None	(256,1)	(256,1)
Conv1D	Kernel size=5 stride = 2	(256,1)	(128,64)
Conv2D	Kernel size=3 stride = 2	(128,64)	(64,128)
Reshape	None	(64,128)	(8192,1)
Max Pooling	pool size=128	(8192,1)	(64,1)
LSTM	units=32	(64,1)	(32,1)
Dense	units=64	(32,1)	(64,1)
Dense	units=32	(64,1)	(32,1)

This paper presents a study related to the use of a modified long and short term memory network for mental workload classification methods. In conducting the model training process, the pre-processed data was used as the input data for the model, where 80% of the data was used to conduct the model training and 20% of the data was used to conduct the test. Details of the parameter settings of the model are shown in Table 3.

The number of times the model was trained was set to 500 and the learning rate was set to 0.001. The experimental results showed that the training accuracy could basically reach over 90% around the 120th epoch. In the first 20 epochs and 70~80 epochs, the discriminant accuracy improved faster, and the experimental result accuracy is shown in Figure 8; the loss value converged more rapidly, and the experimental loss rate is shown in Figure 9. In this model training, the final accuracy rate can reach 94.9%. Compared with other traditional methods, this experimental model has greater accuracy.

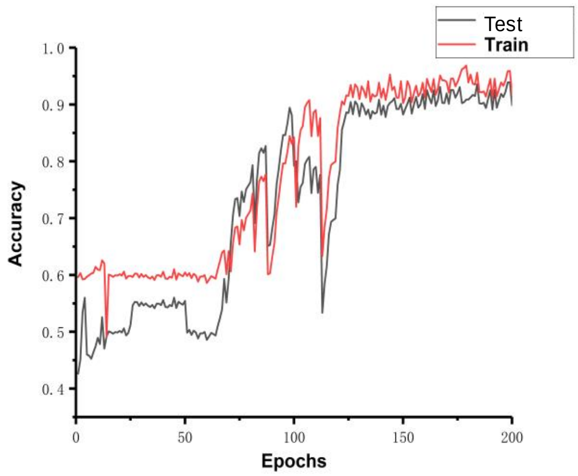


Figure 8. Training and test set accuracy

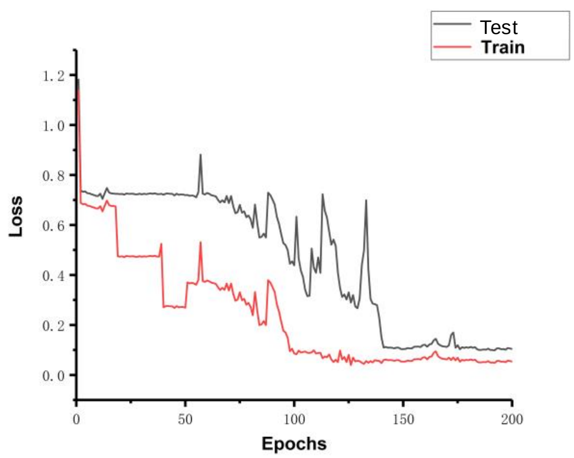


Figure 9. Loss rate for training and test sets

V. CONCLUSIONS

Piloting in aviation is a major hazard and pilots are under great stress during the flight. The study of the pilot's mental workload is conducive to a better assessment of the pilot's driving state, as well as to the effective prediction of phenomena that exceed the mental workload, which is conducive to the rational design of the human-computer interaction system in flight driving, thus reducing the phenomenon of high or low mental workload during flight driving and ensuring the safety of flight driving.

This paper uses a combination of convolutional neural networks and long and short-term memory networks to classify mental workload with an accuracy of 94.9%. The results of the analysis show that there is a positive correlation between the difficulty of the flight task and the subjective evaluation score. In summary, this study has some reference value for the classification of pilot mental workload.

ACKNOWLEDGMENT

This work is supported by The National Natural Science Foundation of China (No. 52072293)

REFERENCES

[1] Previtali Federica, Picco Eleonora, Gragnano Andrea, Miglioretti Massimo. The Relationship between Work, Health and Job Performance for a Sustainable Working Life: A Case Study on Older Manual Employees in an Italian Steel Factory [J]. International Journal of Environmental Research and Public Health, 2022,19(21).

[2] Mohanavelu K, Poonguzhali S, Adalarasu K, Ravi D, Vijayakumar Chinnadurai, Vinutha S, Ramachandran K,Srinivasan Jayaraman. Dynamic cognitive workload assessment for fighter pilots in simulated fighter aircraft environment using EEG [J]. Biomedical Signal Processing and Control, 2020, 61.

[3] Sun Li, Sun Youzhao. Dynamic functional allocation of aircraft cockpit based on load balancing[J]. Measurement and Control Technology, 2022, 41(01):11-15+27.DOI:10.19708/j.ckjs.2021.04.230.

[4] Zhang P, Wang X, Zhang W, et al. Learning spatial–spectral–temporal EEG features with recurrent 3D convolutional neural networks for cross-task mental workload assessment [J]. IEEE Transactions on neural systems and rehabilitation engineering, 2018, 27(1): 31-42.

[5] Mental workload: Its theory and measurement[M]. Springer Science & Business Media, 2013.

[6] Li D, Wang X, Menassa C C, et al. Understanding the impact of building thermal environments on occupants' comfort and mental workload demand through human physiological sensing[M]//Start-Up Creation. Woodhead Publishing, 2020: 291-341.

[7] He Xiaomei, Li Jianyu. Research on the causes analysis and design strategies of cognitive load in human-computer interaction[J]. Packaging Engineering,2020, 41(10):24-30.DOI:10.19554/j.cnki.1001-3563.2020.10.005.

[8] Rainieri G, Fraboni F, Russo G, et al. Visual scanning techniques and mental workload of helicopter pilots during simulated flight[J]. Aerospace Medicine and Human Performance, 2021, 92(1): 11-19.

[9] Longo L. Mental workload in medicine: foundations, applications, open problems, challenges and future perspectives[C]//2016 ieee 29th international symposium on computer-based medical systems (cbms). IEEE, 2016: 106-111.

[10] Dias I M. Work zone crash analysis and modeling to identify the factors affecting crash severity and frequency [M]. Kansas State University, 2015.

[11] Mansikka H, Virtanen K, Harris D. Comparison of NASA-TLX scale, modified Cooper–Harper scale and mean inter-beat interval as measures of pilot mental

workload during simulated flight tasks [J]. Ergonomics, 2019, 62(2): 246-254.

[12] Howell N, Devendorf L, Tian R, et al. Biosignals as social cues: Ambiguity and emotional interpretation in social displays of skin conductance[C]/ /Proceedings of the 2016 ACM Conference on Designing Interactive Systems, Brisbane: ACM, 2016: 865-870.

[13] WY, Yuan XG, Liu ZQ. Analysis of relations between changes of pupil and mental work loads [J]. Space Medicine & Medical Engineering, 2007, 20(5): 364-366.

[14] Hsu CK, Lin SC, Li WC. Visual movement and mental-workload for pilot performance assessment[C]/ /International Conference on Engineering Psychology and Cognitive Ergonomics, Switzerland: Springer International Publishing, 2015: 356- 364.

[15] Liu Yan, Cheng Baoquan, Yu Xin, Gu Haowu, Wang Rui, Xian Jishu, Quan Yulian, Li Fei, Feng Hua, Chen Tunan, Wang Lihua. A typical correlation analysis of brain load, sleep and fatigue in helicopter pilots [J]. Journal of the Army Medical University, 2022, 44(11): 1087-1093. DOI: 10.16016/j. 2097-0927. 202201151.

[16] Wang Lei, Wang Shuo, Zou Ying, Zhang Mengxi, Wu Jianjun, Li Shu. A study on the brain load characteristics of airline pilots based on task context [J/OL]. Journal of Safety and Environment:1-9 [2022-11-20].DOI:10.13637/j.issn.1009-6094.2021.2153.

[17] Wei Zongmin, Wanyan Xiaoru, Zhuang Damin . Measurement and evaluation of mental workload for aircraft cockpit display interface [J]. Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(1) : 86-91.

[18] Grier R A. How high is high? A meta-analysis of NASA-TLX global workload scores[C]//Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Sage CA: Los Angeles, CA: SAGE Publications, 2015, 59(1): 1727-1731.

[19] Chu Hongzuo, Jiao Xuejun, Jiang Jin, Cao Yong, Yang Jiehong, Huang Mengying, Wang Lizhi, Li Xiugan. Characteristics and recognition of personalized EEG response to mental workload [J]. Manned Spaceflight, 2021, 27 (06): 710-718.DOI: 10.16329 / j.cnki. zrht. 2021.06.006.

[20] Chen Qichao, Zhang Xuejun, Huang Wanlu. EMD fusion PSD, CSP EEG feature extraction method [J].Computer technology and development, 2019, 29 (05): 126-130.

[21] Jia Dongwei, Gan Yong, Wang Yifan. Model study based on improved LSTM-CRF combined with attention mechanism [J]. Network security technology and application, 2022 (11): 39-40.